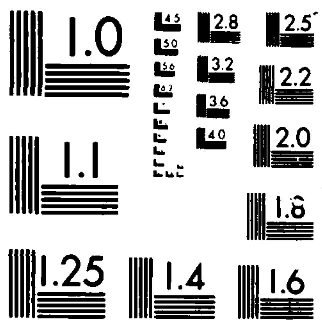
AD-A163 518 ASYMPTOTIC PROPERTIES OF INDUCED MAXIMUM LIKELIHOOD ESTIMATES OF NONLINEA..(U) ADVANCED STATISTICAL TECHNOLOGIES CORP LAWRENCEVILLE NJ D H JONES OCT 85 1/1

UNCLASSIFIED ROB-03-85 N00014-83-C-0627 F/G 12/1 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(12)

AD-A163 518

ASYMPTOTIC PROPERTIES OF INDUCED MAXIMUM LIKELIHOOD ESTIMATES OF NONLINEAR MODELS FOR ITEM RESPONSE VARIABLES: THE FINITE-GENERIC-ITEM-POOL CASE

Douglas H. Jones

Technical Report No. ONR-01-85

Advanced Statistical Technologies Corporation
P.O. Box 6640
Lawrenceville, New Jersey 08648

October 1985

Prepared under contract No. N00014-83-C-0627, NR 150-152
with the Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report No. ONR-01-85	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Asymptotic properties of induced maximum likelihood estimators of nonlinear models for item response variables: the finite-generic-item-pool case		5. TYPE OF REPORT & PERIOD COVERED Technical Report 01 OCT 84 - 31 JUL 85 6. PERFORMING ORG. REPORT NUMBER ROB-03-85
7. AUTHOR(s) Douglas H. Jones		8. CONTRACT OR GRANT NUMBER(s) N00014-83-C-0627
9. PERFORMING ORGANIZATION NAME AND ADDRESS Advanced Statistical Technologies Corporation P.O. Box 6640 Lawrenceville, NJ 08648		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N, RR4204, RR0420401, NR 150-152
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research, Code 442PT Arlington, VA 22217		12. REPORT DATE 01 AUGUST 1985 13. NUMBER OF PAGES 32
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Item Response Theory, Robustness, Ability, Maximum Likelihood, Influence Curve, Jackknife, Breakdown Point, Estimation.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The progress of modern mental test theory depends very much on the techniques of maximum likelihood estimation, and many popular applications make use of likelihoods induced by logistic item response models. While, in reality, item responses are nonreplicate within a single examinee and the logistic models are only ideal, practitioners make inferences using the asymptotic distribution of the maximum likelihood estimator derived as if item responses were replicate and satisfied their ideal model. This article proposes a sample space acknowledging

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

these two realities and derives the asymptotic distribution of the induced maximum likelihood estimator.

This article assumes that items, while sampled from an infinite set of items have but a finite domain of alternate response functions: this situation is the case of the finite-generic-item-pool. Later articles will attempt to remove this assumption.

Using the proposed sample space, the article applies the statistical functional approach of von Mises to derive the influence curve of the maximum likelihood estimator; to discuss related robustness properties; and to derive new classes of resistent estimators. This article's general purpose is revealing the value of these methods for uncovering the relative merits of different item response functions. Proofs and mathematical derivations are minimized to increase the assessability of this complex subject.

Asymptotic properties of induced maximum likelihood estimators of non-linear models for item response variables: the finite-generic-item-pool case

Douglas H. Jones
Advanced Statistical Technologies Corporation

Abstract

The progress of modern mental test theory depends very much on the techniques of maximum likelihood estimation, and many popular applications make use of likelihoods induced by logistic item response models. While, in reality, item responses are nonreplicate within a single examinee and the logistic models are only ideal, practitioners make inferences using the asymptotic distribution of the maximum likelihood estimator derived as if item responses were replicated and satisfied their ideal model. This article proposes a sample space acknowledging these two realities and derives the asymptotic distribution of the induced maximum likelihood estimator.

This article assumes that items, while sampled from an infinite set of items, have but a finite domain of alternate response functions: this situation is the case of the finite-generic-item-pool. Later articles will attempt to remove this assumption.

Using the proposed sample space, the article applies the statistical functional approach of von Mises to derive the influence curve of the maximum likelihood estimator; to discuss related robustness porperties; and to derive new classes of resistent estimators. This article's general purpose is revealing the value of these methods for uncovering the relative merits of different item response functions. Proofs and mathematical derivations are minimized to increase the accessability of this complex subject.

1. INTRODUCTION

While maximum likelihood procedures are popular in item response theory (IRT), (Lord, 1980), their insensitivity to departures from assumptions is serious enough to warrant cautious use and further study (Wainer-Wright, 1980; Jones, 1982). The purpose of this article is to explore the behavior of the procedures when the model is not true.

To apply some of the concepts of robustness theory, we found that some of the more important concepts required reformulating the maximum likelihood procedures. In particular the study of the robustness of the maximum likelihood estimator (MLE) requires viewing it as a function of the empirical probability distribution function (PDF). The original formulation of item response theory, as a regression problem, does not allow the summarization of the data in terms of an empirical PDF. In §2, we recast the structure of the problem so that the data can be replaced by an empirical PDF and we reformulate the MLE as a function of it.

In §3, we derive the asymptotic distribution of the MLE when the true PDF is not generated by the assumed model. These results are basic to understanding the sensitivity of the MLE to departures from assumptions. They make heavy use of von Mises's approach to statistical functions (Fillippova, 1982).

In §4, we apply the asymptotic formulas derived in §3 to three popular item response models. A measure of goodness-of-fit, compatable with the MLE, is employed to play the role of the mean squared error (McCullagh-Nelder, 1983; Pregibon, 1981). These results reveal that certain item response models reverse the scale of ability.

In §5, we formulate the basic robustness criteria associated with Hampel's influence curve (IC) (Hampel, 1974; Welsch-Krasker, 1982). We derive a relation between the IC and the maximum bias of the MLE as the true PDF is varied within an ϵ-contamination neighborhood of the modeled PDF (Huber, 1981). We also derive the breakdown point (Huber, 1981) of the MLE for certain types of departures from the assumptions. The analysis of these criteria shows how the notion of robustness in IRT is fundamentally different from linear and logistic regression problems.

2. GENERAL NOTATION AND STRUCTURE

The basic formulation of IRT based on maximum likelihood is: u=1 (correct) or u=0 (incorrect) is observed for each item i with likelihood, given a real latent parameter θ, equal to

$$h_i(u;\theta) = P_i(\theta)^u \left[1-P_i(\theta)\right]^{1-u}$$

and with $P_i(\theta)$ the i^{th} item response model. The total likelihood based on data $u_1,u_2,\ldots,u_n$ and models $P_1,P_2,\ldots,P_n$ is:

$$L(\theta:u_1,\ldots,u_n,\ P_1,\ldots,P_n) = \prod_{i=1}^{n} h_i(u_i;\theta).$$

For robustness studies, we need to allow for the possibility that the item response models are inaccurate. Thus, we assume that $E(u_i) \neq P_i(\theta)$. But we retain the assumption of local independence and call $P_i(\theta)$ an operational model.

To accommodate items with different difficulties and discriminating powers, and simultaneously, apply standard asymptotic theory, we formulate the sample space as:

$$\text{(Sample Space)}\quad S = \{(u,x):\ u=0,1;\ x\epsilon X\}$$

$$X = \text{finite set indexing items.}$$

An observation on S is denoted by s or t, etc., and is generated by administering a randomly chosen item, x, to obtain a response, u.

An arbitrary probability distribution function (PDF) on S is denoted by η. A probability distribution over X is denoted by p. The conditional probability distribution of u given x is denoted by f(u;x). For arbitrary η, there is a p and f(u;x) such that:

$$\eta(s) = f(u;x)\ p(x),\ s = (u,x).$$

Because u is binary; f(u;x) is Bernoulli with some probability of success, $\Pi^*(x)$ satisfying:

$$f(u;x) = \Pi^*(x)^u\ [1-\Pi^*(x)]^{1-u}.$$

The empirical PDF defined for a sample $s_1,s_2,\ldots,s_n$ is defined by denoting δ_s to be a point mass at s and

$$\hat{\eta}_n(t) = n^{-1} \sum_{i=1}^{n} \delta_{s_i}(t).$$

It is a PDF on S. The distance between two PDF's ξ and η is defined as $|\xi-\eta| = \max_{s\epsilon S}|\xi(s)-\eta(s)|$.

A parametric family of PDF's on S is defined by $\{\eta_\theta:\theta \text{ real}\}$. Values of η_θ are denoted by $\eta(s;\theta)$. A special type of a parametric family is generated by a set of operational models:

$$\text{Operational Models:}\quad \{\Pi(\theta;x):x\epsilon X;\theta \text{ real}\}$$

$$\text{Parametric Family:}\quad f(u;\theta,x)=\Pi(\theta;x)^u\{1-\Pi(\theta;x)\}^{1-u}$$

$$\eta(s;\theta)=f(u;\theta,x)p(x).$$

The traditional structure of IRT is related as follows: the i^{th} observation is u_i with model $P_i(\theta)$. Let x_i be the index value of the i^{th} chosen item where $\Pi(\theta;x_i)=P_i(\theta)$. Let $s_i=(u_i,x_i)$, so that $\eta(s_i;\theta)=f(u_i;\theta,x_i)p(x_i)=h_i(u_i;\theta)$ $p(x_i)$.

The likelihood based on the sample $s_1,s_2,\ldots,s_n$ is:

$$L(\theta;s_1,s_2,\ldots,s_n) = \prod_{i=1}^{n} \eta(s_i;\theta).$$

If $\{p(x_i)\colon i=1,\ldots,n\}$ contains no information about θ, MLE's based on the two likelihoods are identical.

The log-derivative of the parametric PDF is denoted by $\ell(s;\theta) = (d/d\theta) \log \eta(s;\theta)$. If it exists, a solution of the *implicit* equation

$$\text{(Normal Equation)} \qquad 0 = \sum_{i=1}^{n} \ell(s_i;\theta)$$

is denoted by $\hat{\theta}_n$ and is called the MLE. This equation simplifies with operational models as follows: The logit of an operational model $\Pi(\theta;x)$ and its derivative is

$$g(\theta;x) = \log \Pi(\theta;x)/[1-\Pi(\theta;x)]$$

$$g'(\theta;x) = v(\theta;x)^{-1}\ \Pi'(\theta;x)\ \text{, where}$$

$$v(\theta;x) = \Pi(\theta;x)\ [1-\Pi(\theta;x)].$$

Using the definition of η_θ, we have:

$$\ell(s;\theta) = g'(\theta;x)[u-\Pi(\theta;x)]$$

and the normal equation becomes

$$0 = \sum_{i=1}^{n} g'(\theta;x_i)\ [u_i-\Pi(\theta;x_i)] = \sum_{i=1}^{n} v(\theta;x_i)^{-1}\ [u_i-\Pi(\theta;x_i)]\ \Pi'(\theta;x_i)$$

The Fisher information of the parametric PDF, η_θ, is

$$I(\theta) = -\sum \ell'(s;\theta)\ \eta(s;\theta) = \sum \ell(s;\theta)^2\ \eta(s;\theta)$$

where the sum is over all s in S. This information identity follows from the total differential of $0 = \sum \ell(s;\theta)\ \eta(s;\theta)$; using $\eta'(s;\theta) = \eta(s;\theta)(d/d\theta)\log \eta(s;\theta) = \eta(s;\theta)\ \ell(s;\theta)$ we have,

$$0 = \sum \ell'(s;\theta)\eta(s;\theta) + \sum\ell(s;\theta)\eta'(s;\theta)$$

$$= \sum \ell'(s;\theta)\eta(s;\theta) + \sum\ell(s;\theta)^2\eta(s;\theta).$$

Note for computational purposes: $I(\theta) = \sum g'(\theta;x)^2\ v(\theta;x)\ p(x)$.

Example. The one-parameter logistic (1PL) and two-parameter logistic (2PL) item response models are characterized by their logits: $g(\theta;x) = a(x)[\theta-b(x)]$ where $a(x) > 0$, $-\infty < b(x) < \infty$ are the discrimination and difficulty parameters for item x. Hence, $\ell(s;\theta) = a(x)[u-\Pi(\theta;x)]$ and $0 = \sum_{i=1}^{n} a(x_i)[u_i - \Pi(\theta_i;x_i)]$ is the normal equation. The Fisher information is $I(\theta) = \sum a(x)^2\, v(\theta;x)p(x)$, sum over all X.

We wish to generalize the normal equation in two ways: first, we want to show the explicit relation between $\hat{\theta}_n$ and $\hat{\eta}_n$; second, we wish to consider estimators that are more general than MLE's.

We rewrite the normal equation using the empirical PDF as

$$0 = \sum \ell(s;\ \theta)\ \hat{\eta}_n(s)$$

where it will be understood that the sum is always over S. We see that the MLE depends explicitly on the empirical PDF, we denote this dependence by

$$\hat{\theta}_n = \theta(\hat{\eta}_n).$$

If the empirical PDF is replaced by an arbitrary PDF, the normal equation defines a general functional relationship, $\theta(\eta)$, between θ and η: we call $\theta(\eta)$ a statistical functional.

We define M-type estimators generated by a score function $\psi(s;\theta)$ by the equation

$$0 = n^{-1} \sum_{i=1}^{n} \psi(s_i;\ \theta) = \sum \psi(s;\theta)\hat{\eta}_n(s).$$

We see that $\psi(s;\theta)=\ell(s;\theta)$ generates the MLE. We add this generality because our methods of proof in the next section are really about M-type estimators with the MLE results following as a special case. Note that the notion of a statistical function applies to M-type estimators also. More definitions that we need follow.

We define

$$m(\theta,\eta) = \sum \psi(s;\theta)\eta(s)$$

for an arbitrary PDF and score function. The derivative of $m(\theta,\eta)$ with respect to θ is $m'(\theta,\eta)$. Note that $m'(\theta,\eta_\theta)$ means $m'(\theta,\eta)$ evaluated with $\eta=\eta_\theta$. The normal equation is $0=m(\theta,\hat{\eta}_n)$ and Fisher's information is $I(\theta) = -m'(\theta,\eta_\theta)$ with $\psi=\ell$. The Newton-Rapheson algorithm for solving the normal equation is

$$\theta^{t+1} = \theta^t + m(\theta^t,\hat{\eta}_n)/-m'(\theta^t,\hat{\eta}_n);$$

if $\psi=\ell$ the Fisher scoring algorithm is

$$\theta^{t+1} = \theta^t + m(\theta^t,\hat{\eta}_n)/I(\theta^t).$$

Let $\psi(s;\theta)$ be a given score function and let θ_0 denote the value of θ that solves the equation $0 = m(\theta,\eta)$, corresponding to this score function. If the PDF η is a member of some parametric family and satisfies $\eta=\eta_{\theta_1}$ for a given fixed parameter value θ_1 and if $\theta_0=\theta_1$, then we say that the score function is <u>unbiased</u>.

If ψ is an unbiased score function, then $0=m(\theta,\eta_\theta)$ for all θ. This fact leads to an identity that is analogous to the <u>Fisher information</u> identity presented previously and is proven in exactly the same way. The identity is:

$$-m'(\theta,\eta_\theta) = \sum \psi(s;\theta)\ell(s;\theta)\eta(s;\theta).$$

If one replaces $m'(\theta,\hat{\eta}_n)$ by its expectation under η_θ in the Newton-Rapheson algorithm, one obtains an algorithm that is analogous to Fisher scoring. If ψ is an unbiased score function, then one may use the above identity for $-m'(\theta,\eta_\theta)$ to avoid evaluating the derivative of ψ.

An important subclass of unbiased score functions are generated by an arbitrary <u>weight</u> <u>function</u> $w(\theta;x)$ where

$$\psi(s;\theta) = w(\theta;x)\left[u-\Pi(\theta;x)\right]\Pi'(\theta;x).$$

If we choose

$$w(\theta;x) = v(\theta;x)^{-1}$$

then $\psi(s;\theta) = \ell(s;\theta)$ and we are back to the MLE. Other choices of the weight function lead to resistant estimators. For example, Jones (1982) suggests $w(\theta;x) = v(\theta;x)^{h-1}$ with $h \geq 0$, a tuning constant. We can stay in the class of exponential families with arbitrary response variable u, as long as $\Pi(\theta;x)=E(u|\theta,x)$ and $w(\theta;x)=\mathrm{var}(u|\theta,x)^{-1}$ (see Jennrick and Moore, 1975). Jones' resistant estimator could be generated for these families also by letting $w(\theta;x)= \mathrm{var}(u|\theta,x)^{h-1}$. We obtain a more general class of estimators by allowing the weight functions to depend on the response: $\psi(s;\theta)=w(\theta;s)[u-\Pi(\theta;x)]\Pi'(\theta;x)$. Krasker and Welsch (1982) consider these estimators for the general linear model. Stefanski, Carroll, and Ruppert (1984) consider these estimators for the logistic model.

An algorithm based on Gauss-Newton's algorithm for solving normal equations with score $\psi(s;\theta)=w(\theta;s)[u-\Pi(\theta;x)]\Pi'(\theta;x)$ is as follows: (see Holland and Welsch, 1977) define $d_i^t = \Pi'(\theta;x_i)$ and $w_i = w(\theta_i;\ s_i)$ then

$$\theta^{t+1} = \theta^t + [\sum d_i^t\ w_i^t\ d_i^t]^{-1} \sum d_i^t\ w_i^t\ [u_i-\Pi(\theta^t;x_i)].$$

This algorithm is iterative reweighted least squares: at convergence $\hat{\theta}_n=\theta^\infty$, so if we define the pseudo-observation $z_i=d_i\hat{\theta}_n + [u_i - \Pi(\hat{\theta}_n;x_i)]$ then

$$\hat{\theta}_n = [\sum d_i\ w_i\ d_i]^{-1} \sum d_i\ w_i\ z_i.$$

See Pregibon (1981) and McCullagh and Nelder (1983) for a similar algorithm based on the exponential family with linear predictors. Note that this algorithm is also identical to Fisher scoring.

3. GENERAL ASYMPOTOTIC THEORY OF M-TYPE ESTIMATORS

We present consistency and asymptotic normality (AN) results in this section. In the first part we confine attention to the main results and in the second part we supply the proofs. Readers may skip the proofs and move on to the next section. In the main results we discuss conditions for consistency and AN. We also characterize an approximation to the M-type estimator that is important for AN results and for the robustness results in §5.

3.1 Main Results: Consistency

Suppose the sample $s_1, s_2, \ldots, s_n$ is IID with PDF η. The empirical PDF $\hat{\eta}_n$ satisfies $|\hat{\eta}_n - \eta| \to 0$ wp 1 as $n \to \infty$. This fact gives us the obvious candidate for the limit of an M-type estimator, θ_0 which solves $0 = m(\theta_0, \eta)$; when does $\hat{\theta}_n \to \theta_0$? It is possible that the equation $0 = m(\theta, \hat{\eta}_n)$ yielding $\hat{\theta}_n$ has more than one solution, in which case a consistency result may be about only one of the possible sequences of M-type estimators. Also, it is possible that the equation $0 = m(\theta_0, \eta)$ does not have a local solution, in which case a consistency result would not be useful. Some known results follow.

Huber (1964): Let θ_0 be the unique solution. If $\psi(s;\theta)$ is monotone in θ for each $s \epsilon S$ then <u>every</u> sequence $\hat{\theta}_n \to \theta_0$ wp1.

Boos (1977): Let θ_0 be an isolated solution. If $\psi(s,\theta)$ is continuous in θ for each $s \epsilon S$ then <u>there exists</u> a sequence $\hat{\theta}_n \to \theta_0$ wp1.

Huber (1967, 1980): Let θ_0 be an unique solution. Let $|m(\theta,\eta)|$ be bounded from zero as $|\theta| \to \infty$. If $\psi(s;\theta)$ is continuous in θ for each $s \epsilon S$, then <u>every</u> sequence $\hat{\theta}_n \to \theta_0$ wp1.

The various conditions for consistency will be satisfied when we impose stricter conditions for AN.

Example. For the 2PL model $\Pi(\theta;x)$ is strictly monotone increasing and hence

$\psi(s;\theta) = a(x)[u-\Pi(\theta;x)]$ is monotone, the solution θ_0 to $0=m(\theta_0,\eta)=$ $\Sigma\, a(x)[\Pi^*(x)-\Pi(\theta_0;x)]p(x)$ is unique with $\Pi^*(x)$ arbitrary. Thus for η arbitrary, Huber (1964) applies. If $\eta=\eta_{\theta_1}$ for some fixed θ_1, then $\theta_0=\theta_1$ is the unique solution.

3.2 Main Results: Asymptotic Normality

The first order asymptotic properties of M-type estimators are characterized by the influence curve. The influence curve (IC) is defined as: let $s\epsilon S$ and δ_s a point mass at s, for $\epsilon>0$

$$IC(s,\eta,\psi) = \lim_{\epsilon\to 0} \frac{\theta(\eta+\epsilon(\delta_s-\eta))-\theta(\eta)}{\epsilon} .$$

Denoting $\theta(\delta_s;\epsilon) = \theta(\eta+\epsilon(\delta_s-\eta))$, we see that the influence curve is an ordinary derivative of $\theta(\delta_s;\epsilon)$ evaluated at 0: $IC(s,\eta,\psi) = (d\theta(\delta_s;\epsilon)/d\epsilon)_0$. For M-type estimators, it will be proved later that for ξ an arbitrary PDF, $\sum IC(s,\eta,\psi)\xi(s) = (d\theta(\xi;\epsilon)/d\epsilon)_0$. This latter characterization allows us in §5 to make an important connection between the bias and the influence curve of MLE's. Also for M-type estimators the influence curve is:

(**) $$IC(s,\eta,\psi) = \psi(s;\theta_0)/-m'(\theta_0,\eta)$$

where throughout the remainder of this section θ_0 satisfies $0=m(\theta_0,\eta)$ and $m'(\theta_0,\eta)<0$.

Example. For the 2PL and η arbitrary

$$IC(s,\eta,\psi) = a(x)[u-\Pi(\theta_0;x)]/\sum a(x)^2\, v(\theta_0;x)p(x).$$

Normally the denominator would depend explictly on $\Pi^*(x)$ but does not, since $g''(\theta;x)=0$. However, it does depend on $\Pi^*(x)$ through θ_0 which solves $0 = \sum a(x)[\Pi^*(x)-\Pi(x;\theta_0)]p(x)$.

The primary application we make of the influence curve in this section is to get a leading term approximation:

(*) $$\hat{\theta}_n-\theta_0 = n^{-1} \sum_{i=1}^{n} IC(s_i,\eta,\psi)+R_n$$

where R_n is the remainder term. We show below $n^{\frac{1}{2}} R_n \to 0$ in probability, thus the behavior of $n^{\frac{1}{2}}(\hat{\theta}_n-\theta_0)$ is deduced from the approximation and the Lindeberg-Levy central limit theorem. The sufficient conditions for AN of M-type estimators are listed under (C). They are implied by conditions (D) when the M-type estimator is an MLE:

(C) There exists an open interval Ω_0 and a constant $c>0$ such that for all θ in Ω_0

C-1: $\psi(s;\theta)$; $\psi'(s;\theta)$; $\psi''(s;\theta)$ exist for all $s\epsilon S$, with the first two continuous in θ;

C-2: $m'(\theta_\epsilon,\eta_\epsilon) < 0$ for all $0 \le \epsilon \le 1$ and ξ: $|\eta-\xi|\le\epsilon$ where θ_ϵ solves $0=m(\theta,\eta_\epsilon)$ and $\eta_\epsilon = \eta + \epsilon(\xi-\eta)$.

(D) Define $\Omega_0 = \{\theta\colon v(\theta;x)>0 \text{ for all } x\epsilon X\}$ then suppose Ω_0 is not empty and for all $\theta\epsilon\Omega_0$

D-1: $\Pi'(\theta;x)$; $\Pi''(\theta;x)$; $\Pi'''(\theta;x)$ exist for all x, with the first two continuous in θ;

D-2: $\sum g''(\theta;x)[\Pi^*(x)-\Pi(\theta;x)]p(x) < \sum g'(\theta;x)^2 v(\theta;x)p(x)$ for $\Pi^*(x)$ in an open interval for each x.

Theorem 1. Assume (C) then $n^{\frac{1}{2}}(\hat{\theta}_n-\theta_0)$ is AN mean 0 and variance:

$$\sigma_0{}^2 = \sum IC(s,\eta,\psi)^2 \eta(s).$$

Corollary. Assume (C-2) and $\eta=\eta_{\theta_1}$, θ_1 fixed. If ψ is unbiased, i.e. $\theta_0=\theta_1$, $n^{\frac{1}{2}}[\hat{\theta}_n-\theta_0]$ is AN with mean 0 and variance:

$\sigma_0^2 = \sum \psi(s;\theta_0)^2\eta(s;\theta_0)/[\sum \psi(s;\theta_0)\ell(s;\theta_0)\eta(s;\theta_0)]^2$ where $\ell(s;\theta) = (\partial/\partial\theta) \log \eta(s;\theta)$. Hence an unbiased M-type estimator is efficient if and only if ψ is proportional to ℓ, or in other words the MLE is optimal among M-type estimators with unbiased score functions.

3.3 Proof of Asymptotic Normality

Define for any ξ PDF,

$$\eta_\epsilon = \eta + \epsilon(\xi-\eta) \quad \text{and}$$

$$\theta_\epsilon = \theta(\eta_\epsilon).$$

A second order Taylor series expansion about 0 is

$$(***) \qquad \theta_\epsilon = \theta_0 + (d\theta/d\epsilon)_0\epsilon + \tfrac{1}{2}\,(d^2\theta/d\epsilon^2)_a\ a^2,$$

$0 < a < \epsilon$. Under conditions (C) we show that this expansion is valid with $\epsilon=1$. Using the expansion with $\xi=\hat{\eta}_n$, we obtain the approximation (*), where R_n is expressed in terms of the second derivative of θ_ϵ. To show that the IC has the form (**) and that $n^{\frac{1}{2}} R_n \to 0$ in probability we derive the first two derivatives as follows.

Define $M(\theta,\epsilon) = m(\theta,\eta_\epsilon)$. Because θ_ϵ satisfies $M(\theta_\epsilon,\epsilon) = 0$ for all $0\le\epsilon\le1$, the first and second total differentials with respect to ϵ are identically zero and yield two simultaneous equations involving $d^j\theta/d\epsilon^j$ j=1,2:

$$(1) \qquad (\partial M/\partial\theta)_{\theta_\epsilon}\ d\theta/d\epsilon + \partial M/\partial\epsilon = 0$$

$$(2) \qquad (\partial M/\partial\theta)_{\theta_\epsilon}\ d^2\theta/d\epsilon^2 + [(\partial^2 M/\partial\theta^2)_{\theta_\epsilon}\ (d\theta/d\epsilon) + (\partial^2 M/\partial\theta\partial\epsilon)_{\theta_\epsilon}]d\theta/d\epsilon$$
$$+ (\partial^2 M/\partial\epsilon\partial\theta)_{\theta_\epsilon}\ d\theta/d\epsilon + \partial^2 M/\partial\epsilon^2 = 0.$$

Solving equations (1) and (2) for $d^j\theta/d\epsilon^j$, j=1,2, and using $\partial M^2/\partial\epsilon^2 = 0$, we have from equation (1):

$$(1') \qquad \frac{d\theta}{d\epsilon} = \left(\frac{\partial M}{\partial\epsilon}\right) \Big/ - \left(\frac{\partial M}{\partial\theta}\right)_{\theta_\epsilon}$$

and from equations (1') and (2):

$$(2') \qquad \frac{d^2\theta}{d\epsilon^2} = \left(\frac{\partial^2 M}{\partial\theta^2}\right)_{\theta_\epsilon} \left(\frac{\partial M}{\partial\epsilon}\right)^2 \Big/ - \left(\frac{\partial M}{\partial\theta}\right)^3_{\theta_\epsilon} + 2\left(\frac{\partial^2 M}{\partial\epsilon\,\partial\theta}\right)_{\theta_\epsilon} \left(\frac{\partial M}{\partial\epsilon}\right) \Big/ \left(\frac{\partial M}{\partial\theta}\right)^2_{\theta_\epsilon}$$

Now we obtain expression (**) for the IC: Let $\theta_0 = \theta(\eta)$, we have $\partial M/\partial\epsilon = (\partial/\partial\epsilon)\ m(\theta_0,\eta_\epsilon) = (\partial/\partial\epsilon) \sum \psi(s;\theta_0)[\eta(s) + \epsilon(\xi(s) - \eta(s))]$ $= \sum \psi(s;\theta_0)\ [\xi(s) - \eta(s)] = \sum \psi(s;\theta_0)\ \xi(s)$. Substituting $(\partial M/\partial\theta)_{\theta_0} = m'(\theta_0,\eta)$

and the last expression into (1') we have for any PDF ξ, $(d\theta/d\epsilon)_0 = \sum \psi(s;\theta_0)$ $\xi(s)/ -m'(\theta_0,\eta)$. Thus for $\xi = \delta_s$, we have $IC(s,\eta,\psi) = \psi(s;\theta_0)/ -m'(\theta_0,\eta)$.

With $\xi=\hat{\eta}_n$, the empirical PDF, the MLE $\hat{\theta}_n = \theta(\eta_1)$ where η_1 is η_ϵ with $\epsilon=1$. Using the expansion (***) with $\epsilon=1$ we have

$$\hat{\theta}_n - \theta_0 = n^{-1} \sum_{i=1}^{n} IC(s_i,\eta,\psi) + R_n$$

where

$$R_n = \tfrac{1}{2}\left(\frac{d^2\theta}{d\epsilon^2}\right)_{a^*} \text{, for some } 0<a^*<1.$$

Now we state the conditions for the expansion (***) and hence the expression for R_n to be valid: (see Serfling, 1980, pp. 43, 215).

(A) Apostle (1957, pg. 96) $(d\theta/d\epsilon)^+$, the righthand derivative, and $d^2\theta/d\epsilon^2$ exist everywhere in the open interval (0,1); with the first continuous in the half-closed interval [0,1).

By expression (1') and (2') for $d\theta/d\epsilon$ and $d^2\theta/d\epsilon^2$ we have formulated conditions (B) that satisfy conditions (A):

(B) There exists an open interval Ω_0 such that for all θ in Ω_0

B-1: $m(\theta,\eta)$, $m'(\theta,\eta)$, $m''(\theta,\eta)$ exist for all η;

B-2: there exists a constant c, such that for all ξ: $|\xi-\eta|\le c$, $m'(\theta_\epsilon,\eta_\epsilon)<0$ for all $0\le\epsilon\le1$.

To further obtain $n^{\frac{1}{2}}R_n \to 0$, we need to examine the terms in expression (2') and place appropriate conditions on the score function, ψ. The four terms are:

$$(\partial M/\partial\theta)_{\theta_\epsilon} = m'(\theta_\epsilon,\eta_\epsilon)$$

$$(\partial^2 M/\partial\theta^2)_{\theta_\epsilon} = m''(\theta_\epsilon,\eta)$$

$$(\partial M/\partial\epsilon) = m(\theta_\epsilon,\xi-\eta)$$

$$(\partial M/\partial\epsilon\,\partial\theta)_{\theta_\epsilon} = m'(\theta_\epsilon,\xi-\eta).$$

These terms apply to R_n with $\xi=\hat{\eta}_n$.

We see that the behavior of R_n depends directly on that of $\hat{\eta}_n$ and in particular the differences $\hat{\eta}_n(s) - \eta(s)$, $s\epsilon S$. We have given condition (C) at the beginning of this section to keep $m'(\theta_\epsilon,\eta_\epsilon)$ properly away from zero so as to keep R_n from exploding and to infer its behavior from that of $\hat{\eta}_n$. The following two lemmas imply that $n^{\frac{1}{2}} R_n \to 0$ wp1.

Lemma A. Assume conditions (C). Put $\xi=\hat{\eta}_n$ in (2'). There exists constants a and n_0 such that $|d^2\theta/d\epsilon^2| \le a|\hat{\eta}_n-\eta|^2$ for all $0\le\epsilon\le 1$ and $n\ge n_0$ wp1.

Lemma B. Assume $s_1,s_2,\ldots,s_n$ are IID with PDF η. Then as $n\to\infty$ the following holds:

a) $\hat{\eta}_n(s)\to\eta(s)$ for all $s\epsilon S$ wp1;

b) $\{n^{\frac{1}{2}}[\hat{\eta}_n(s) - \eta(s)]: s\epsilon S\}$ converges in law to a Gaussian process with mean 0 and covariance function:

$$\mathrm{COV}(s,t) = \begin{cases} \eta(s)[1-\eta(s)] & s=t \\ \\ -\eta(s)\eta(t) & s\neq t\ ; \end{cases}$$

c) $|\hat{\eta}_n-\eta|\to 0$ wp1;

d) $n^{\frac{1}{2}}|\hat{\eta}_n-\eta|$ converges in law and in probability.

These two lemmas imply that $n^{\frac{1}{2}} R_n\to 0$ in probability and hence $n^{\frac{1}{2}}(\hat{\theta}_n-\theta_0)$ is AN since lemma A implies $n^{\frac{1}{2}}|R_n|\le a\ n^{\frac{1}{2}}|\eta_n-\eta|\cdot|\eta_n-\eta|$ and lemma B implies that $|\eta_n-\eta|\to 0$ while $n^{\frac{1}{2}}|\hat{\eta}_n-\eta|$ remains bounded in probability.

4. SPECIFIC ASYMPTOTIC BEHAVIOR OF MLE's

Throughout this section η_0 will denote a true PDF and θ_0 the solution to $0=m(\theta,\eta_0)$. The asymptotic behavior of the MLE is obtained from the previous results of M-type estimators with score function $\ell(s;\theta)$, $s\epsilon S$. We discuss these

aspects: goodness-of-fit, scale reproduction, and Fisher variance.

Most practical operational models satisfy condition (D-1). Thus, we can say that for situations of interest, MLE's are consistent, $\hat{\theta}_n \to \theta_0$ as $n \to \infty$ and is asymptotically normal even when η_0 is not a member of the parametric family $\{\eta_\theta : \theta \text{ real}\}$ generated by the set of operational models but satisfies the mild regularity condition (D-2). We will employ this result with the 1PL, 2PL, and 3PL item response models in the example at the end of this section.

Let η_θ denote a member of a certain parametric family and consider the MLE associated with the family. If for some θ_1, $\eta_0 = \eta_{\theta_1}$, then the MLE is asymptotically unbiased, meaning $\theta_0 = \theta_1$. Let ξ_{θ_1} denote a member of a different parametric family. If $\eta_0 = \xi_{\theta_1}$ then the MLE is asymptotically biased meaning $\theta_0 \neq \theta_1$. If η_0 is not a member of any parametric family then the notion of unbiasedness has no meaning.

Even if η_0 is a member of some parametric family not identical to $\{\eta_\theta\}$, the bias does hold much information about how good the MLE may be. This is because the parametrization of the family containing η_0 is as good as arbitrary when it is not exactly the one generating the MLE. This leads us to propose a different notion of accuracy, possibly supplying the information we usually obtain with measurements of bias.

The information supplied by the bias is obtained from comparison of its square to the variance, because the mean square error, a measure of total error, is the sum of the squared bias and variance. When the bias overwhelms the variance, one usually goes looking for another statistical procedure that can control the bias. (What this compares to in IRT is the adoption of more complex item response models).

A measure that seems to decompose into parts due to "bias" and "variance" and does not depend on the arbitrary parametrization of a true family of PDF's is as follows: for two PDF's η and ξ define

$$K(\eta,\xi) = \sum \eta(s) \log \eta(s)/\xi(s).$$

$K(\eta,\xi)$ is nonnegative and equal to 0 when $\eta=\xi$; $K'(\eta_0,\eta_\theta) = -m(\theta,\eta_0)$, thus $K(\eta_0,\eta_\theta)$ is minimized by θ_0. A second order Taylor series expansion gives

$$E\ K(\eta_0,\eta_{\hat{\theta}_n}) \cong K(\eta_0,\eta_{\theta_0}) - \tfrac{1}{2}\ m'(\theta_0,\eta_0)n^{-1}\sigma_0^2,$$

where we have assumed that the MLE has been modified appropriately to have the moments for the approximation and we have used Theorem 1.

Thus $E\ K(\eta_0,\eta_{\hat{\theta}_n})$ behaves as a "total error" and $K(\eta_0,\eta_{\theta_0})$ behaves as "bias squared" when it is compared to the last term on the right-hand side of the above Taylor series. The quantity $K(\hat{\eta}_n,\eta_{\hat{\theta}_n})$ is proportional to the deviance in generalized linear models (see McCullagh and Nelder, 1983) and serves as a goodness-of-fit statistic. Values of $E\ K(\eta_0,\eta_{\hat{\theta}_n})$ and its approximate components are displayed in Table 3 and discussed in the example at the the end of this section.

Information of a different nature than bias, applicable to arbitrarily parametrized families, is obtained by comparing the rank-order of estimated parameters with the rank-order of known abilities. Suppose there is a certain parametric family $\{\xi_\theta\}$ of PDF's with the property that $\eta_0 \in \{\xi_\theta\}$ where η_0 may be generated by any member of a population of examinees. But the family $\{\xi_\theta\}$ is too complex, making its calibration unstable with reasonable sample sizes of examinees. Thus, we prefer instead to use a more parsimonious family $\{\eta_\theta\}$ with the MLE obtaining θ_0 as a limit when $\eta_0=\xi_{\theta_1}$, θ_1 fixed, and $n\to\infty$. Previous discussion implies that $\theta_0\neq\theta_1$, in general; but the bias here is nonsense. What

is useful is a measure of the distortion between θ_0 and θ_1 as θ_1 moves throughout the population. We do not propose a measure but we do think that one should be sensitive to reversals in the "θ-scale". Table 3 displays reversals for the 1PL and 2PL families and is further discussed in the example at the end of this section.

Turning from bias to variance, we will now consider the predicament of approximating the true asymptotic variance of an MLE when we do not know η_0. A ready approximation to σ_0^2 of Theorem 1 in §3 is the reciprocal of Fisher's information: $I(\theta_0)^{-1}$ (§2). But we know from §3 that it is not valid when η_0 is not a member of the parametric family of PDFs that generate the MLE.

In general, the σ_0^2 does not majorize $I(\theta_0)^{-1}$ or viceversa. Thus, it is possible that the reciprocal of Fisher's information can give either a conservative or misleading approximation to the true asymptotic variance. Table 3 displays the true variance along side the Fisher variance to show both the good and the bad; we discuss this further in the example.

Example. Listed in Table 1 are the true and modeled response probabilities of five subjects on four ASVAB items. The true probabilities were actually obtained from a very complex item response model which was calibrated on a very large population. The subjects are ranked from lowest to highest going left to right. The modeled response probabilities follow the 1PL, 2PL, or 3PL item response models as indicated; they too were calibrated on a very large population; Table 2 lists the values of the calibrated parameters. Table 3 is a summary of the asymptotic features of the respective MLE's, which we discuss as follows.

First, note the magnitudes of the traditional notion of bias by taking $|\theta_1-\theta_0|$ differences from columns (1) and (2). These values could easily change

upon reparametrization of the true response model; thus they are arbitrary. Thus column (1) should only convey the rank-order of the subjects.

Our notion of "bias squared" is found in column (5), $K(\eta_0,\theta_0)$. These values will not change if another parametrization were imposed on the true model. The worst fit is found with subject 5(2PL), referring back to Table 1 we can see that the 2PL model provides poor estimates of all item response probabilities. There are three good values; for example, subject 5(3PL) for which Table 1 shows good estimates of item response probabilities.

Column (3), $-m'(\theta_0,\eta_0)$, gives us a feel for the curvature of the likelihood since $n \cdot m'(\theta_0,\eta_0)$ is an estimate of $n \cdot m'(\hat{\theta}_n,\hat{\eta}_n)$, the second derivative of the log-likelihood. We see that the likelihood would tend to be flat for subjects 1,2 (3PL) even though $K(\eta_0,\theta_0)$ shows close agreement between the estimated and true item response probabilities.

We present the "total error" $E\,K(\eta_0,\eta_{\hat{\theta}_n})$, column (4), for a sample size of n=16, each item type represented equally. These errors appear to be equal across models and subjects with exception of subject 5 (2PL) as noted before. The components of the total error are in columns (5) and (6) which can tell us the proportion of the total error due to systematic bias: (5)/(4). The worst proportion is found with subjects 1, 2 (1PL) meaning that the 1PL is inadequate with these subjects.

We may average the "total error" and "bias squared," $E\,K(\eta_0,\eta_{\theta_n})$ and $K(\eta_0,\eta_{\theta_0})$ respectively, over the subjects to get an overall assessment. These averages are for the 1PL, 2PL, and 3PL models respectively: (error, bias^2) = (.06,.03), (.06,.03), (.04,.01). We see that on average there is at least 25 percent of total error that is systematic bias.

The presence of reversals of the θ_0-scale can be detected from column (2). Both the 1PL and 2PL item response models have reversals at the lower abilities. This happens because the 1PL and 2PL calibrations compensate good fit to true response probabilities by distorting the θ_0-scale. The Spearman rank correlation between the true ability rank-order and the θ_0-scale rank-order of the 1PL, 2PL, and 3PL models are respectively: 0.60, 0.67, 1.00. The numbers may be interpreted as an alternative <u>theoretical</u> goodness-of-fit, since we never would know the true rank-order of ability, there is no practical gain in the measure.

We compare the true variance and the Fisher variance by using columns (7) and (8). For the most part, the Fisher variance yields a conservative assessment of precision; however, it can also be misleading as with subjects 3, 5 (2PL).

Remarks. 1) Column (3), $-m'(\theta_0,\eta_0)$, can play the role of information. An empirical assessment could be $-m'(\hat{\theta}_n,\hat{\eta}_n)$. Also, ratios could play the role of relative efficiency.

2) One should be cautious even if measures of fit, such as $K(\hat{\eta}_n,\eta_{\hat{\theta}_n})$, are favorable because as the example shows it is possible to have reversals of the θ_0-scale even if the fit is good.

3) We have refrained from making an elaborate comparison of the 1PL, 2PL, and 3PL models based on the data, because one needs to properly account for sampling variability of the calibration process. Such a study is reported in Jones, Wainer and Kaplan (1984).

5. SPECIFIC ROBUSTNESS OF THE MLE

Let η denote an arbitrary true PDF, η_0 some fixed PDF, $\{\eta_\theta\}$ a parametric family of PDF's that induces the MLE. Let $\theta(\eta)$ denote the solution to $0=m(\theta,\eta)$.

From Theorem 1 in §3 we have that $\theta_n \to \theta(\eta)$ with asymptotic variance $\sigma^2 = \sigma^2(\eta)$. Note that we will not assume that η or η_0 belongs to $\{\eta_\theta\}$.

The asymptotic bias of the MLE relative to η_0 and $\{\eta_\theta\}$ is defined as $|\theta(\eta) - \theta(\eta_0)|$. Let P_ϵ denote an ϵ-neighborhood of η_0, for η belonging to P_ϵ we want to quantify the degradation of bias and variance. We say that the robustness of the MLE is measured by the amount of degradation of the maximum bias

$$b(\epsilon) = \sup_{\eta \epsilon P_\epsilon} |\theta(\eta) - \theta(\eta_0)|$$

and the maximum variance

$$v(\epsilon) = \sup_{\eta \epsilon P_\epsilon} \sigma^2(\eta).$$

If $b(\epsilon)$ were large relative to $v(\epsilon)$, then the maximum variance would not be a very important quantifier of robustness. We confine study to $b(\epsilon)$ in this paper.

There are several important notions for quantifying the robustness of an estimator. Among them are the sensitivities of a parameteric estimator, a fitted value, or a predicted value when one observation is deleted from the sample. These measures are called, respectively, gross error sensitivity (Huber, 1981), change in fit sensitivity and prediction sensitivity (Krasker and Welsch, 1983). The gross error sensitivity is related directly to the maximum bias as shown below. We formulate these quantities and demonstrate their use with the 1PL, 2PL, and 3PL item response models.

Another robustness notion is the sensitivity of the maximum bias as ϵ is varied. Certain values of ϵ can cause the maximum bias to explode; the smallest such value is called the breakdown point (Huber, 1981). We formulate this quantity and demonstrate its use with the 1PL, 2PL, and 3PL models also.

5.1 Sensitivities Based on Deletion

The <u>gross error sensitivity</u> is defined as

$$\gamma^* = \max_s |IC(s,\eta_0,\psi)|.$$

From the leading term approximation of section 3.2, we see that it is proportional to the maximal influence exerted by anyone observation on the error of estimation, $\hat{\theta}_n - \theta(\eta_0)$. It is related to the maximum bias $b(\epsilon)$ as follows. Recall from §3.2 that $[\theta(\eta_0+\epsilon(\xi-\eta_0))-\theta(\eta_0)] \;\epsilon\rightarrow \sum IC(s,\eta_0,\psi)\xi(s)$ as $\epsilon\rightarrow 0$. Let P_ϵ be the <u>ϵ-contamination neighborhood</u> defined by $P_\epsilon = \{\eta\colon \eta=\eta_0 + \epsilon(\xi-\eta_0),\ \xi$ arbitrary PDF$\}$. Then

$$\sup_{\eta\in P_\epsilon}|\theta(\eta)-\theta(\eta_0)| \cong \epsilon \sup_\xi |\sum IC(s,\eta_0,\psi)\xi(s)|.$$

Thus

$$b(\epsilon) \cong \epsilon\ \gamma^*.$$

So that for small ϵ, γ^* measures the rate of growth of the maximum bias over the ϵ-contaminated neighborhood.

For M-type estimators $\gamma^* = \infty$ is equivalent to a zero breakdown point, meaning that any departure from η_0 will cause the maximum bias to explode. Either condition also implies that the estimator is not continuous at η_0 when reviewed as a function of η (assuming, of course, a complimentary topology on the set of PDF's). An estimator is qualitatively robust if it is continuous (Huber, 1981), thus an M-type estimator is <u>not</u> robust if $\gamma^* = \infty$ or the breakdown point is zero.

The gross error sensitivity also measures the maximum change in the estimator caused by deleting one observation. Let $\hat{\eta}_n(1)$ and $\hat{\eta}_n(0)$ denote the empirical PDF with and without s_i. Let $\hat{\theta}_n(1)$ and $\hat{\theta}_n(0)$ denote the corresponding MLE's. Then using the direct definition of the influence curve (§3.2) with $s=s_i$,

$\eta=\hat{\eta}_n(0)$ and $\epsilon=1/n$ it is easy to show

$$\theta_n(1) - \theta_n(0) \cong n^{-1}\ IC(s_i,\eta_n(0),\psi).$$

Thus

$$\max_i |\hat{\theta}_n(1) - \hat{\theta}_n(0)| \cong n^{-1}\ \gamma^*.$$

The change in fit sensitivity concerns the effect of deleting one observation, s_i, on the estimated logit, $g(\hat{\theta}_n;x_i)$. This change in fit is $g(\hat{\theta}_n(1);x_i) - g(\hat{\theta}_n(0);x_i) \cong g'(\hat{\theta}_n(0);x_i)[\hat{\theta}_n(1) - \hat{\theta}_n(0)]$. Putting this together with the estimator sensitivity we have

$$g(\hat{\theta}_n(1);x_i)-g(\hat{\theta}_n(0);x_i)\cong n^{-1}g'(\hat{\theta}_n(0);x_i)IC(s_i,\hat{\eta}_n(0),\psi).$$

Thus the shape of $g'(\theta;x)IC(s,\eta,\psi)$ would indicate robustness as would the size of the <u>change in fit sensitivity</u>:

$$\gamma^{**} = \max_s|g'(\theta;x)IC(s,\eta,\psi)|.$$

Prediction sensitivity concerns the effect of deleting an observation from the sample on the predicted logit of ne future item, $g(\hat{\theta}_n;z)$ where z is yet to be administered. Let $\lambda=g'(\theta;z)$, then by a Taylor series approximation, $g(\hat{\theta}_n;z)\cong g(\theta;z)+\lambda(\hat{\theta}_n-\theta)$. Hence the change in prediction is measured by the change in $\lambda\hat{\theta}_n$, and $\lambda IC(s_i,\eta,\psi)$ measures this change due to deleting s_i. To be meaningful this change must be weighed relative to its standard deviation, $\lambda[\sum IC(s,\eta,\psi)^2\ \eta(s)]^{\frac{1}{2}}$. Thus the shape of the ratio indicates robustness as would the <u>prediction sensitivity</u>:

$$\gamma = \max_s \frac{|IC(s,\eta,\psi)|}{[\sum IC(s,\eta,\psi)^2\ \eta(s)]^{\frac{1}{2}}}.$$

We can simplify this quantity to show the direct dependence on the score function by using the formula for the influence curve:

$$\gamma = \max_s \frac{|\psi(s;\theta)|}{[\sum \psi(s;\theta)^2 \; \eta(s)]^{\frac{1}{2}}}$$

where θ is evaluated at $\theta(\eta)$.

Now we study the various sensitivities to get a feel for their implications in IRT using the 1PL, 2PL, and 3PL models as examples. Graphs of these quantities are useful but require specific values for item parameters and do not lead to any more profound conclusions then just analytic circumspection. Graphs are most useful, however, with actual data, providing diagnostic information on the fit of the model. We study only the MLE induced by $\{\eta_\theta\}$ and do not look at general M-type estimators. We also restrict this study to sensitivities to departures from the parametric model, that is we let $\eta_0=\eta_\theta$ for some value of θ. Huber (1980) remarks that a better indication of robustness is to allow η to roam around a P_ϵ neighborhood of η_θ while looking at the sensitivities. We do not have the analytical means to do this at this time.

Consider now and for the rest of this section the MLE with operational models $\{\Pi(\theta;x):x\epsilon X\}$. With $\eta_0=\eta_\theta$, $-m'(\theta;\eta_\theta) = \sum \psi(s;\theta)\; \ell(s;\theta)\; \eta(s;\theta)$ and with $\psi(s;\theta) = \ell(s;\theta) = g'(\theta;x)[u-\Pi(\theta;x)]$, we have

$$IC(s,\eta,\ell) = \frac{g'(\theta;x)[u-\Pi(\theta;x)]}{\sum g'(\theta;x)^2 \; v(\theta;x)\; p(x)} .$$

Define

$$M(\theta;x) = \max\{\Pi(\theta;x),\; 1-\Pi(\theta;x)\},$$

the various sensitivies to departures from $\eta_0=\eta_\theta$ are

$$\gamma^* = \frac{\max_x g'(\theta;x)\; M(\theta;x)}{\sum g'(\theta;x)^2 \; v(\theta;x)\; p(x)} .$$

$$\gamma^{**} = \frac{\max_x g'(\theta;x)^2\; M(\theta;x)}{\sum g'(\theta;x)^2 \; v(\theta;x)\; p(x)} \quad \text{and}$$

$$\gamma = \frac{\max\limits_{x} g'(\theta;x)\ M(\theta;x)}{\left[\sum g'(\theta;x)^2\ v(\theta;x)\ p(x)\right]^{\frac{1}{2}}} .$$

Example 1. The 2PL item response models have $g'(\theta;x) = a(x)$ where $\infty > a(x) > 0$. If $a(x) \equiv a_0$, then the models are called the 1PL item response models. γ^* is finite provided max a(x) is finite. If the generic item pool X is finite then γ^* is always finite. If the generic item pool is not finite, it is possible that $\sup a(x) = \infty$ but practical reasons would disallow this from happening because an "infinitely discriminating item" is rare.

Example 2. The 3PL item response models are defined as $\Pi(\theta;x) = [1-c(x)]\ R(\theta;x) + c(x)$ where $0<c(x)<1$ and $R(\theta;x)$ is a 2PL model. Define $v_1(\theta;x) = R(\theta;x)[1-R(\theta;x)]$. It can be shown that $g'(\theta;x) = [1-c(x)]a(x)\ v_1(\theta;x)/v(\theta;x)$. γ^* is finite provided max a(x) is finite, the discussion in the previous example applies here too.

Example 3. For all the 1PL, 2PL, and 3PL models, because of the behavior of $g'(\theta;x)$, γ^{**} and γ are finite if and only if max a(x) is finite. γ^{**} and γ, but not γ^*, are invariant for changes of scale in a(x) and b(x). Presumably c(x) is scale free as it is a probability of the examinee guessing the correct answer to item x.

The examples lead to the general conclusion that γ, γ^*, and γ^{**} are finite if and only if $\max|g'(\theta;x)|$ is finite. For the 1PL, 2PL, and 3PL models this condition is equivalent to having max a(x) finite.

Because the sensitivities change as θ changes, their variation over the entire range of practical θ-values should be studied to properly assess robustness in IRT. This allows for the fact that the MLE procedure must estimate unique θ parameters for different subjects. This is in marked contrast with estimation in logistic regression -- the same estimation procedures as IRT but

the object is to estimate a single θ (such as lethal dose 50 or the vector of parameters in one response function). Because the sensitivities must be viewed globally, procedures that are robust for logistic regression may not be directly transferable to IRT.

Consider what happens as $|\theta|\to\infty$. The denominator of γ^* and γ^{**} is Fisher's information; for γ, it is just the square root. For extreme θ's it is reasonable to assume that any finite set of generic items X, item responses hold little information about θ; thus it is probable that the denominators of the sensitivities approach zero as $|\theta|\to\infty$. Unless the numerators approach zero at the same or faster rate as the denominators, the sensitivities will explode. Applying this idea to each sensitivity, we conclude that γ^* always explodes and for models with $v(\theta;x)\to 0$, γ^* and γ^{**} both explode. Of the models considered before, the 3PL is the only one having $v(\theta;x)>0$ as $\theta\to-\infty$; thus γ^* and γ^{**} are bounded for the negative extremes of ability.

These results imply that the MLE procedures are not robust because the maximum bias in an ϵ-contaminated neighborhood is approximately $\epsilon\gamma^*$ and γ^* is unbounded as $|\theta|\to\infty$; thus, the MLE cannot tolerate any contamination at extreme θ. The 3PL fairs a little better than the 1PL or 2PL as $\theta\to-\infty$ since its gross error sensitivity grows a little slower. Thus to achieve full protection one must look outside the class of MLE procedures, which means we have to sacrifice efficiency. (Contrast this with the location problem where the median is the efficient procedure for logistic errors and it is optimal for minimizing the maximum bias; Huber, 1981).

5.2 Breakdown Point

The worst possible bias at η_0 is defined as $b(1) = \sup_{\xi}|\theta(\xi)-\theta(\eta_0)|$, where the supremum is over all arbitrary PDFs, ξ. Let P_ϵ be an ϵ-neighborhood of η_0.

The <u>breakdown point</u>, ϵ^*, is the largest ϵ for which $b(\epsilon)$ is less than the worst value:

$$\epsilon^* = \sup\{\epsilon : b(\epsilon) < b(1)\}.$$

The value of ϵ^* depends on the kind of P_ϵ chosen; however, it is sometimes adequate to consider just one kind of neighborhood. In IRT, $b(1)=\infty$.

We use the following kind of ϵ-neighborhood: Let $0 \le \Pi \le 1$ and define $v = \Pi(1-\Pi)$. Denote the interval $D(\theta;x) = [\Pi - \epsilon v^{\frac{1}{2}}, \Pi + \epsilon v^{\frac{1}{2}}]$ when $\Pi = \Pi(\theta,x)$; θ is fixed. Denote the subinterval of $[0,1]$ by $D^*(\theta;x) = D(\theta;x) \cap [0,1]$. The collection of intervals $\{D^*(\theta;x)\}$, x fixed, is an ϵ-envelope of the item response function $\Pi(\theta;x)$. Define $P_{\epsilon,\theta} = \{\eta : \eta(s) = \Pi^*(x)^u\ [1-\Pi^*(x)]^{1-u}\ p(x);\ \Pi^*(x) \epsilon\ D^*(\theta;x)\}$. It is an ϵ-neighborhood "centered" at η_θ.

Define $b_+(\epsilon) = \sup_\xi[\theta(\xi)-\theta(\eta_0)]$ and $b_-(\epsilon) = \inf_\xi[\theta(\xi)-\theta(\eta_0)]$. Then $b(\epsilon) = \max\{b_+(\epsilon), - b_-(\epsilon)\}$. We consider $b_+(\epsilon)$ first.

Let $\eta_0 = \eta_{\theta_0}$. Define $\Pi^+(\theta_0;x) = \min\{1, \Pi+\epsilon v^{\frac{1}{2}}\}$ with $\Pi=\Pi(\theta_0;x)$. It is clear that $\Pi^+(\theta_0;x)\ \epsilon\ D(\theta_0;x)$ and $\overset{+}{\eta}_{\theta_0}$, the corresponding PDF, satisfies $m(\theta;\overset{+}{\eta}_{\theta_0}) > m(\theta;\eta)$ for all θ and all $\eta\ \epsilon\ P_{\epsilon,\theta_0}$. The maximum "positive" bias satisfies

$$b_+(\epsilon) = \inf\{\theta : m(\theta;\overset{+}{\eta}_{\theta_0}) < 0\} - \theta_0.$$

We have breakdown if $b_+(\epsilon) = b(1) = \infty$. To avoid this it is necessary that ϵ satisfy $\lim_{\theta\to\infty} m(\theta;\overset{+}{\eta}_{\theta_0}) < 0$. Using the definition of $m(\theta;\eta)$ we have

$$m(\theta;\eta_{\theta_0}) = \sum g'(\theta;x)[\Pi(\theta_0;x) - \Pi(\theta;x)]p(x) + \epsilon \sum g'(\theta;x)v(\theta_0;x)^{\frac{1}{2}}p(x).$$

Letting $\theta\to\infty$ and denoting $g'(\infty;x) = \lim g'(\theta;x)$ we have an equation for the "positive" side breakdown:

$$\epsilon^+ = \frac{\sum g'(\infty;x)\ [1-\Pi(\theta_0;x)]\ p(x)}{\sum g'(\infty;x)\ v(\theta_0;x)^{\frac{1}{2}}\ p(x)} .$$

Similarly the "negative" side breakdown is:

$$\epsilon^- = \frac{\sum g'(-\infty;x)\ \Pi(\theta_0;x)\ p(x)}{\sum g'(-\infty;x)\ v(\theta_0;x)^{\frac{1}{2}}\ p(x)}\ .$$

And the breakdown:

$$\epsilon^+ = \min(\epsilon^+, \epsilon^-).$$

For a fixed θ_0, all MLE procedures for IRT have a breakdown point that is not 0 for P_ϵ neighborhoods considered thus far. But as $|\theta_0|\to\infty$ the picture changes: $\epsilon^* = 0$ if either $\Pi(\theta_0;x)\to 1$ or 0 for all items x. Thus the 1PL and 2PL induced MLEs have zero breakdown, meaning they have no tolerance for departures from their models. The 3PL induced MLE has zero breakdown, but for $\theta\to-\infty$, the "negative" sided breakdown is not zero, so it could tolerate some departure from its model there.

Example. The following displays the "positive" and "negative" breakdown points for the 3PL model with $a(x) = a_0$ and $c(x) = c_0$.

c_0	ϵ^-	ϵ^+
.025	.16	0
.05	.23	0
.10	.33	0
.20	.50	0

REFERENCES

Apostle, T.M. (1957). Mathematical Analysis. Reading, MA.: Addison-Wesley.

Boos, D.D. (1977). The differential approach in statistical theory and robust inference. Tallahassee, FL: Florida State University, Ph.D. Dissertation.

Flippova, A.A. (1962). Mises' theorem of the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. Theor. Prob. App., 7, 24-57.

Hampel, F.R. (1968). Contributions to the theory of robust estimation. Berkeley, CA: University of California, Ph.D. Dissertation.

Hampel, F.R. (1974). The influence curve and its role in robust estimation. Journal of the American Statistical Association, 62, 1179-1186.

Holland, P.W. and Welsch, R.E. (1977). Robust regression using iteratively reweighted least-squares. Communication in Statistics: Theory and Methods, 6, 813-827.

Huber, P.J. (1964). Robust estimation of a location parameter. Annals of Mathematical Statistics, 35, 73-101.

Huber, P.J. (1967). The behavior of maximum likelihood estimates with non-standard conditions. In Proceedings of the Fifth Berkeley Symposium on Mathematics Statistics and Probability, Volume 1. Berkeley: University of California Press.

Huber, P.J. (1981). Robust Statistics. New York: Wiley and Sons.

Jennrich, R.I. and Moore, R.H. (1975). Maximum likelihood estimation by means of nonlinear least squares. Proceedings of the Statistical Computing Section, 57-65. Washington, DC: American Statistical Association.

Jones, D.H. (1982). Redescending M-type estimators of latent ability. Princeton, NJ: Educational Testing Service, Program Statistics Research Technical Report No. 82-30 and Research Report No. 82-15.

Jones, D.H., Wainer, H., and Kaplan, B. (1984). Estimating ability with three item response models when the models are wrong and their parameters are inaccurate. Princeton, NJ: Educational Testing Service, Program Statistics Research Technical Report No. 84-46 and Research Report No. 84-26.

Krasker, W.S. and Welsch, R.E. (1982). Efficient bounded-influence regression estimation. Journal of the American Statistical Association, 77, 595-604.

Lord, F.M. (1980). Application of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Erlbaum Associates.

McCullagh, P. and Nelder, P.A. (1983). Generalized Linear Models. New York: Chapman and Hall.

Pregibon, D. (1981). Logistic regression diagnostics. Annals of Statistics, 9, 705-724.

Serfling, R.J. (1980). Approximation Theorems of Mathematical Statistics. New York: Wiley and Sons.

Stefanski, L.A., Carroll, R.J., and Ruppert, D. (1984). Bounding influence and leverage in logistic regression. Ithaca, NY: Cornell University, Department of Economic and Social Statistics.

Wainer, H. and Wright, B.D. (1980). Robust estimation of ability in the Rasch model. Psychometrika, 45, 373-390.

TABLE 1. ITEM RESPONSE PROBABILITIES

MODEL	ITEM 1	ITEM 2	ITEM 3	ITEM 4
Subject:1				
True	.42	.27	.03	.36
1PL	.32	.30	.19	.32
2PL	.30	.30	.08	.30
3PL	.26	.25	.04	.23
Subject:2				
True	.32	.26	.05	.26
1PL	.26	.25	.15	.10
2PL	.27	.27	.05	.26
3PL	.26	.26	.05	.23
Subject:3				
True	.19	.28	.11	.20
1PL	.22	.21	.12	.07
2PL	.27	.27	.05	.26
2PL	.27	.28	.10	.24
Subject:4				
True	.50	.44	.47	.60
1PL	.55	.53	.38	.55
2PL	.46	.46	.50	.53
3PL	.46	.43	.46	.41
Subject:5				
True	.89	.77	.88	.95
1PL	.88	.88	.79	.88
2PL	.72	.72	.98	.84
3PL	.91	.73	.85	.96

TABLE 2. PARAMETERS OF OPERATIONAL MODELS $\{\eta_\theta\}$

1-PL

Item	b	a	c
1	.9	.7	0
2	1.0	.7	0
3	1.9	.7	0
4	.9	.7	0

2-PL

Item	b	a	c
1	1.4	.5	0
2	1.4	.5	0
3	1.1	1.7	0
4	.9	.7	0

3-PL

Item	b	a	c
1	1.3	3.2	.26
2	1.6	1.9	.25
3	1.1	2.1	.03
4	1.1	3.5	.23

TABLE 3. ASYMPTOTIC PARAMETERS OF MLE's

	(1)	(2)†	(3)	(4)*	(5)	(6)*	(7)	(8)
Model	θ_1	θ_0	$-m'(\theta_0,\eta_0)$	$EK(\eta_0,\eta_{\hat{\theta}_0})$	$K(\eta_0,\eta_{\theta_0})$	$-\frac{1}{2}m'(\theta_0,\eta_0)n^{-1}\sigma_0^2$	σ_0^2	$I(\theta_0)^{-1}$
Subject:1								
1PL	-2	-.2(3)	.10	.07	.04	.03	9.47	10.26
2PL	-2	-.3(3)	.10	.04	.02	.02	7.61	9.76
3PL	-2	-1.2(1)	.00+	.03	.03	.00+	400.00	400.00
Subject:2								
1PL	-1	-.6(2)	.10	.07	.04	.03	8.68	10.26
2PL	-1	-.6(1.5)	.09	.03	.00+	.03	11.42	11.76
3PL	-1	-.8(2)	.01	.04	.00+	.04	133.33	133.33
Subject:3								
1PL	0	-.9(1)	.08	.06	.03	.03	13.33	13.33
2PL	0	-.6(1.5)	.09	.06	.02	.04	15.92	11.76
3PL	0	-.1(3)	.06	.04	.01	.03	18.06	16.67
Subject:4								
1PL	1	1.2(4)	.12	.04	.01	.03	8.16	8.16
2PL	1	1.1(4)	.24	.05	.02	.03	4.12	4.12
3PL	1	1.0(4)	.96	.05	.02	.03	1.03	1.04
Subject:5								
1PL	2	3.8(5)	.06	.06	.03	.03	14.08	16.00
2PL	2	3.3(5)	.06	.14	.07	.07	36.36	18.18
3PL	2	1.9(5)	.60	.03	.00+	.03	1.68	1.66

* n=16

† Rank order appears in parenthesis.

Department of Defense

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Dr. Anita Lancaster
Accession Policy
OASD/MI&L/MP&FM/AP
Pentagon Room 2B271
Washington, DC 20301

1 Dr. Jerry Lehnus
OASD (M&RA)
Washington , DC 20301

1 Dr. Clarence McCormick
HQ, MEPCOM
MEPCT-P
2500 Green Bay Road
North Chicago, IL 60064

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 Dr. W. Steve Sellman
Office of the Assistant Secretary
of Defense (MRA & L)
2B269 The Pentagon
Washington, DC 20301

1 Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria , VA 22333

Civilian Agencies

1 Dr. Patricia A. Butler
NIE-BRN Bldg, Stop # 7
1200 19th St., NW
Washington, DC 20208

1 Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E Street NW
Washington, DC 20415

1 Mr. Thomas A. Warm
U. S. Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Private Sector

1 Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

1 Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

1 Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
Israel

1 Dr. Werner Birke
Personalstammamt der Bundeswehr
D-5000 Koeln 90
WEST GERMANY

1 Dr. R. Darrell Bock
Department of Education
University of Chicago
Chicago, IL 60637

1 Mr. Arnold Bohrer
Section of Psychological Research
Caserne Petits Chateau
CRS
1000 Brussels
Belgium

1 Dr. Robert Brennan
American College Testing Programs
P. O. Box 168
Iowa City, IA 52243

1 Dr. Glenn Bryan
6208 Poe Road
Bethesda, MD 20817

1 Dr. Ernest R. Cadotte
307 Stokely
University of Tennessee
Knoxville, TN 37916

1 Dr. John B. Carroll
409 Elliott Rd.
Chapel Hill, NC 27514

Private Sector

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. Hans Crombag
Education Research Center
University of Leyden
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

1 Lee Cronbach
16 Laburnum Road
Atherton, CA 94205

1 CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

1 Mr. Timothy Davey
University of Illinois
Department of Educational Psychology
Urbana, IL 61801

1 Dr. Dattprasad Divgi
Syracuse University
Department of Psychology
Syracuse, NY 13210

1 Dr. Emmanuel Donchin
Department of Psychology
University of Illinois
Champaign, IL 61820

1 Dr. Hei-Ki Dong
Ball Foundation
800 Roosevelt Road
Building C, Suite 206
Glen Ellyn, IL 60137

1 Dr. Fritz Drasgow
Department of Psychology
University of Illinois
603 E. Daniel St.
Champaign, IL 61820

1 Dr. Stephen Dunbar
Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

Private Sector

1 Dr. John M. Eddins
University of Illinois
252 Engineering Research Laboratory
103 South Mathews Street
Urbana, IL 61801

1 Dr. Susan Embertson
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
Lawrence, KS 66045

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Benjamin A. Fairbank, Jr.
Performance Metrics, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

1 Dr. Leonard Feldt
Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

1 Univ. Prof. Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

1 Professor Donald Fitzgerald
University of New England
Armidale, New South Wales 2351
AUSTRALIA

1 Dr. Dexter Fletcher
University of Oregon
Department of Computer Science
Eugene, OR 97403

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01002

Private Sector

1 Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

1 Dr. James G. Greeno
University of California, Berkeley
Department of Education

Berkeley , CA

1 Dipl. Pad. Michael W. Habon
Universitat Dusseldorf
Erziehungswissenshaftliches Inst. II
Universitatsstr. 1
D-4000 Dusseldorf 1
WEST GERMANY

1 Dr. Ron Hambleton
School of Education
University of Massachusetts
Amherst, MA 01002

1 Prof. Lutz F. Hornke
Universitat Dusseldorf
Erziehungswissenschaftliches Inst. II
Universitatsstr. 1
Dusseldorf 1
WEST GERMANY

1 Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

1 Dr. Lloyd Humphreys
Department of Psychology
University of Illinois
603 East Daniel Street
Champaign, IL 61820

Navy

1 Code N711
Attn: Arthur S. Blaiwes
Naval Training Equipment Center
Orlando, FL 32813

1 Dr. Nick Bond
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

1 Lt. Alexander Bory
Applied Psychology
Measurement Division
NAMRL
NAS Pensacola, FL 32508

1 Dr. Robert Breaux
NAVTRAEQUIPCEN
Code N-095R
Orlando, FL 32813

1 Dr. Robert Carroll
NAVOP 115
Washington, DC 20370

1 Dr. Stanley Collyer
Office of Naval Technology
800 N. Quincy Street
Arlington, VA 22217

1 CDR Mike Curran
Office of Naval Research
800 N. Quincy St.
Code 270
Arlington, VA 22217

1 Dr. John Ellis
Navy Personnel R&D Center
San Diego, CA 92252

1 DR. PAT FEDERICO
Code P13
NPRDC
San Diego, CA 92152

1 Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152

1 Ms. Rebecca Hetter
Navy Personnel R&D Center (Code 62)
San Diego, CA 92152

Navy

1 Mr. Dick Hoshaw
NAVOP-135
Arlington Annex
Room 2834
Washington, DC 20350

1 Dr. Norman J. Kerr
Chief of Naval Education and Training
Code 00A2
Naval Air Station
Pensacola, FL 32508

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Daryll Lang
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. William L. Maloy (02)
Chief of Naval Education and Training
Naval Air Station
Pensacola, FL 32508

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr William Montague
NPRDC Code 13
San Diego, CA 92152

1 Ms. Kathleen Moreno
Navy Personnel R&D Center (Code 62)
San Diego, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Personnel & Training Research Program
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Navy

1 Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152

1 Dr. Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Alfred F. Smode
Senior Scientist
Code 7B
Naval Training Equipment Center
Orlando, FL 32813

1 Dr. Richard Snow
Liaison Scientist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Edward Wegman
Office of Naval Research (Code 411S&P)
800 North Quincy Street
Arlington, VA 22217

1 Dr. Ronald Weitzman
Naval Postgraduate School
Department of Administrative
Sciences
Monterey, CA 93940

1 Dr. Douglas Wetzel
Code 12
Navy Personnel R&D Center
San Diego, CA 92152

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Navy

1 Mr John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Marine Corps

1 Col. Ray Laidich
Headquarters, Marine Corps
MPI
Washington, DC 20380

1 Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 Major Frank Yohannan, USMC
Headquarters, Marine Corps
(Code MPI-20)
Washington, DC 20380

Army

1 Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

1 Dr. Clessen Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

1 Dr. Karen Mitchell
Army Research Institute
5001 Eisenhower Blvd
Alexandria, VA 22333

1 Dr. William E. Nordbrock
FMC-ADCO Box 25
APO, NY 09710

1 Dr. Harold F. O'Neil, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Commander, U.S. Army Research Institute
for the Behavioral & Social Sciences
ATTN: PERI-BR (Dr. Judith Orasanu)
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Mr. Robert Ross
U.S. Army Research Institute for the
Social and Behavioral Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joyce Shields
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Air Force

1 Dr. William E. Alley
AFHRL/MOT
Brooks AFB
, TX 78235

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

1 Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

1 Dr. Sherrie Gott
AFHRL/MODJ
Brooks AFB
, TX 78235

1 Dr. Patrick Kyllonen
AFHRL/MOE
Brooks AFB, TX 78235

1 Dr. Randolph Park
AFHRL/MOAN
Brooks AFB, TX 78235

1 Dr. Roger Pennell
Air Force Human Resources Laboratory
Lowry AFB, CO 80230

1 Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

1 Maj. Bill Strickland
AF/MPXOA
4E168 Pentagon
Washington, DC 20330

1 Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

1 Major John Welsh
AFHRL/MOAN
Brooks AFB , TX 78223

Private Sector

1 Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

1 Dr. Jack Hunter
2122 Coolidge St.
Lansing, MI 48906

1 Dr. Huynh Huynh
College of Education
University of South Carolina
Columbia, SC 29208

1 Dr. Douglas H. Jones
Advanced Statistical Technologies
Corporation
10 Trafalgar Court
Lawrenceville, NJ 08148

1 Professor John A. Keats
Department of Psychology
The University of Newcastle
N.S.W. 2308
AUSTRALIA

1 Dr. William Koch
University of Texas-Austin
Measurement and Evaluation Center
Austin, TX 78702

1 Dr. Thomas Leonard
University of Wisconsin
Department of Statistics
1210 West Dayton Street
Madison, WI 53705

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Private Sector

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
Netherlands

1 Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

1 Dr. Robert Lockman
Center for Naval Analysis
200 North Beauregard St.
Alexandria, VA 22311

1 Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

1 Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

1 Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

1 Mr. Robert McKinley
University of Toledo
Dept of Educational Psychology
Toledo, OH 43606

1 Dr. Barbara Means
Human Resources Research Organization
300 North Washington
Alexandria, VA 22314

1 Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

1 Dr. W. Alan Nicewander
University of Oklahoma
Department of Psychology
Oklahoma City, OK 73069

Private Sector

1 Dr. Melvin R. Novick
356 Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

1 Dr. James Olson
WICAT, Inc.
1875 South State Street
Orem, UT 84057

1 Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Cirle, NW
Washington, DC 20036

1 Dr. James Paulson
Dept. of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

1 Dr. Lawrence Rudner
403 Elm Avenue
Takoma Park, MD 20012

1 Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 Frank L. Schmidt
Department of Psychology
Bldg. GG
George Washington University
Washington, DC 20052

Private Sector

1 Lowell Schoer
Psychological & Quantitative
Foundations
College of Education
University of Iowa
Iowa City, IA 52242

1 Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaigan
Fujusawa 251
JAPAN

1 Dr. William Sims
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. H. Wallace Sinaiko
Program Director
Manpower Research and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

1 Dr. Paul Speckman
University of Missouri-Columbia
Department of Statistics
Columbia, MO 65201

1 Martha Stocking
Educational Testing Service
Princeton, NJ 08541

1 Dr. Peter Stoloff
Center for Naval Analysis
200 North Beauregard Street
Alexandria, VA 22311

1 Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

1 Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Lab
252 Engineering Research Laboratory
Urbana, IL 61801

Private Sector

1 Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

1 Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

1 Mr. Gary Thomasson
University of Illinois
Department of Educational Psychology
Champaign, IL 61820

1 Dr. Robert Tsutakawa
Department of Statistics
University of Missouri
Columbia, MO 65201

1 Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

1 Dr. V. R. R. Uppuluri
Union Carbide Corporation
Nuclear Division
P. O. Box Y
Oak Ridge, TN 37830

1 Dr. David Vale
Assessment Systems Corporation
2233 University Avenue
Suite 310
St. Paul, MN 55114

1 Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08540

1 Dr. Ming-Mei Wang
Lindquist Center for Measurement
University of Iowa
Iowa City, IA 52242

1 Dr. Brian Waters
HumRRO
300 North Washington
Alexandria, VA 22314

Private Sector

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 Dr. Rand R. Wilcox
University of Southern California
Department of Psychology
Los Angeles, CA 90007

1 German Military Representative
ATTN: Wolfgang Wildegrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washington, DC 20016

1 Dr. Bruce Williams
Department of Educational Psychology
University of Illinois
Urbana, IL 61801

1 Ms. Marilyn Wingersky
Educational Testing Service
Princeton, NJ 08541

1 Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, NY 10021

1 Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940

END

FILMED

3-86

DTIC